STIC-Biotech/ChemLib 93134

From: Stucker, Jeffrey
Sent: Monday, May 05, 2003 10:36 AM
To: STIC-Biotech/ChemLib
Subject: 09/868399

Please search SEQ ID NO 1. The peptide is limited to no more that 60 amino acids in the claims.

Thanks,
Jeff Stucker
1648
308-4237
mail: 8E12

	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Searcher:_____	NA Sequences:_____	STN:_____
Phone:_____	AA Sequences:_____	DIALOG:_____
Location:_____	Structures:_____	Questel/Orbit:_____
Date Picked Up:_____	Bibliographic:_____	DRLink:_____
Date Completed: 5/7/03	Litigation:_____	Lexis/Nexis:_____
Searcher Prep/Review:_____	Full text:_____	Sequence Sys.:_____
Clerical:_____	Patent Family:_____	WWW/Internet:_____
Online time:_____	Other:_____	Other (specify):_____

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:55:59 ; Search time 29 Seconds
(without alignments)
255.783 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 124338

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147	74.2	90	15	Q88124	Q88124 chimpanzee
2	147	74.2	90	15	Q88130	Q88130 chimpanzee
3	112	56.6	93	15	Q87646	Q87646 chimpanzee
4	56	28.3	59	15	Q9Q071	Q9q071 chimpanzee
5	53.5	27.0	70	16	P70950	P70950 bacillus su
6	51	25.8	96	6	Q9BE76	Q9be76 macaca fasc
7	50.5	25.5	91	16	Q97KS3	Q97ks3 clostridium
8	50	25.3	44	15	Q71895	Q71895 human immun
9	50	25.3	44	15	Q71902	Q71902 human immun
10	50	25.3	44	15	Q71909	Q71909 human immun
11	50	25.3	44	15	Q71916	Q71916 human immun
12	50	25.3	44	15	Q71923	Q71923 human immun
13	50	25.3	44	15	Q71929	Q71929 human immun
14	50	25.3	44	15	Q71936	Q71936 human immun
15	50	25.3	44	15	Q71942	Q71942 human immun
16	50	25.3	44	15	Q71874	Q71874 human immun
17	50	25.3	44	15	Q71879	Q71879 human immun
18	50	25.3	44	15	Q71885	Q71885 human immun
19	50	25.3	44	15	Q71889	Q71889 human immun
20	50	25.3	44	15	Q71967	Q71967 human immun
21	50	25.3	44	15	Q71972	Q71972 human immun
22	50	25.3	44	15	Q71978	Q71978 human immun
23	50	25.3	44	15	Q71986	Q71986 human immun
24	50	25.3	44	15	Q71997	Q71997 human immun
25	50	25.3	44	15	Q72004	Q72004 human immun
26	50	25.3	44	15	Q72012	Q72012 human immun
27	50	25.3	57	5	Q9TX09	Q9tx09 geodia cydo
28	49	24.7	77	12	Q8V2L8	Q8v218 camelpox vi
29	48.5	24.5	95	16	Q9K0G0	Q9k0g0 neisseria m
30	45	22.7	44	15	Q72008	Q72008 human immun
31	45	22.7	44	15	Q72018	Q72018 human immun
32	45	22.7	65	15	Q04098	Q04098 simian t-ce
33	45	22.7	82	16	Q92XC8	Q92xc8 rhizobium m
34	44	22.2	44	15	Q71991	Q71991 human immun
35	44	22.2	80	4	O15521	O15521 homo sapien
36	43.5	22.0	47	4	Q96CJ4	Q96cj4 homo sapien
37	43	21.7	52	4	Q9UHS7	Q9uhs7 homo sapien
38	43	21.7	62	4	Q96FS0	Q96fs0 homo sapien
39	43	21.7	77	13	Q90XF4	Q90xf4 coturnix co
40	42.5	21.5	48	11	Q925T2	Q925t2 mus musculu
41	42	21.2	60	16	Q98CP2	Q98cp2 rhizobium l
42	42	21.2	76	16	Q8ZSC0	Q8zsc0 anabaena sp
43	41.5	21.0	96	16	Q9I6C5	Q9i6c5 pseudomonas
44	41	20.7	94	16	O31639	O31639 bacillus su
45	40.5	20.5	38	4	Q9BXG9	Q9bxg9 homo sapien

ALIGNMENTS

RESULT 1
Q88124
ID Q88124 PRELIMINARY; PRT; 90 AA.
AC Q88124;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pol polyprotein (Fragment).
GN GAG-POL OR POL.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90272009; PubMed=1971917;
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
RT "Sequence analysis and acute pathogenicity of molecularly cloned
RT SIV.";
RL Nature 345:636-640(1990).
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC (BY SIMILARITY).
DR EMBL; L03296; AAA47761.1; -.
DR InterPro; IPR001037; Integrase_C.
DR Pfam; PF00552; integrase; 1.
KW Endonuclease; Hydrolase; Nucleotidyltransferase; Polyprotein;
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
SQ SEQUENCE 90 AA; 10238 MW; 3FA32AB86D4B57FF CRC64;

Query Match 74.2%; Score 147; DB 15; Length 90;
Best Local Similarity 70.6%; Pred. No. 1.1e-13;
Matches 24; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQ 34
           | :|||||||: || ||||| :|||||||||:::
Db      16 KFKNFRVYYREGRDQLWKGPGELLWKGEGAVILK 49
```

RESULT 2

```
Q88130
ID   Q88130      PRELIMINARY;      PRT;    90 AA.
AC   Q88130;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pol polyprotein (Fragment).
GN   GAG-POL OR POL.
OS   Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11723;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90272009; PubMed=1971917;
RA   Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
RT   "Sequence analysis and acute pathogenicity of molecularly cloned
RT   SIV.";
RL   Nature 345:636-640(1990).
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC       (BY SIMILARITY).
DR   EMBL; L03297; AAA47769.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
KW   Endonuclease; Hydrolase; Nucleotidyltransferase; Polyprotein;
KW   RNA-directed DNA polymerase.
FT   NON_TER       1       1
SQ   SEQUENCE   90 AA;  10238 MW;  3FA32AB86D4B57FF CRC64;

  Query Match             74.2%;  Score 147;  DB 15;  Length 90;
  Best Local Similarity   70.6%;  Pred. No. 1.1e-13;
  Matches   24;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQ 34
           | :|||||||: || ||||| :|||||||||:::
Db      16 KFKNFRVYYREGRDQLWKGPGELLWKGEGAVILK 49


RESULT 3
Q87646
ID   Q87646      PRELIMINARY;      PRT;    93 AA.
AC   Q87646;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pol polyprotein (Fragment).
GN   POL.
OS   Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11723;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Janssens W., Fransen K., Peeters M., Heyndrickx L., Motte J.,
RA   Bedjabaga L., Delaporte E., Piot P., van der Groen G.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yesley R.;
RL   Submitted (JUN-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U11495; AAA61759.1; -.
DR   InterPro; IPR001584; Rve.
DR   Pfam; PF00665; rve; 1.
KW   Polyprotein.
FT   NON_TER       1       1
FT   NON_TER      93      93
SQ   SEQUENCE   93 AA;  10628 MW;  63069A60AFD75771 CRC64;

  Query Match             56.6%;  Score 112;  DB 15;  Length 93;
  Best Local Similarity   83.3%;  Pred. No. 1.3e-08;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYRDSRDPLWKGPAKLL 24
           ||| |||||| ||||:||||| ||
Db      70 KIQKFRVYYRGSRDPIWKGPASLL 93


RESULT 4
Q9Q071
ID   Q9Q071      PRELIMINARY;      PRT;    59 AA.
AC   Q9Q071;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Pol polyprotein (Fragment).
GN   POL.
OS   Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11723;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20057949; PubMed=10590144;
RA   Corbet S., Muller-Trutwin M.C., Versmisse P., Delarue S., Ayouba A.,
RA   Lewis J., Brunak S., Martin P., Brun-Vezinet F., Simon F.,
RA   Barre-Sinoussi F., Mauclere P.;
RT   "env Sequences of Simian Immunodeficiency Viruses from Chimpanzees in
RT   Cameroon Are Strongly Related to Those of Human Immunodeficiency Virus
RT   Group N from the Same Geographic Area.";
RL   J. Virol. 74:529-534(2000).
DR   EMBL; AF115394; AAF18582.1; -.
KW   Polyprotein.
FT   NON_TER       1       1
FT   NON_TER      59      59
SQ   SEQUENCE   59 AA;  6719 MW;  30D5461462B338BC CRC64;

  Query Match             28.3%;  Score 56;  DB 15;  Length 59;
  Best Local Similarity   83.3%;  Pred. No. 0.99;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYRDS 12
           |:||| ||||||
Db      48 KVQNFPVYYRDS 59


RESULT 5
P70950
ID   P70950      PRELIMINARY;      PRT;    70 AA.
AC   P70950; O08143;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein yitx.
GN   YITX.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W168 CA;
RX   MEDLINE=97177785; PubMed=9025291;
RA   Levine A., Vannier F., Roche B., Autret S., Mavel D., Seror S.J.;
RT   "A 10.3 kbp segment from nprB to argJ at the 102 degrees region of the
RT   Bacillus subtilis chromosome.";
RL   Microbiology 143:175-177(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98015415; PubMed=9353931;
RA   Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT   "Sequencing of regions downstream of addA (98 degrees) and citG (289
RT   degrees) in Bacillus subtilis.";
RL   Microbiology 143:3305-3308(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
```

```
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
DR   EMBL; Z79580; CAB01839.1; -.
DR   EMBL; Y09476; CAA70635.1; -.
DR   EMBL; Z99109; CAB12957.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   70 AA;  7757 MW;  A3490E71E2CDDE66 CRC64;

  Query Match             27.0%;  Score 53.5;  DB 16;  Length 70;
  Best Local Similarity   42.9%;  Pred. No. 2.8;
  Matches   12;  Conservative    5;  Mismatches   10;  Indels    1;  Gaps    1;

Qy       6 RVYYRDSRDPLWKGPAKLLWKGEGAVVI 33
           |: ||  | |  |   |::|| | ||::
Db      44 RLLYR-RRPPCRKTGKKIIWKNEDAVMV 70


RESULT 6
Q9BE76
ID   Q9BE76      PRELIMINARY;      PRT;    96 AA.
AC   Q9BE76;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Hypothetical 10.7 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN PARIETAL LOBE;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB056770; BAB39318.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   96 AA;  10672 MW;  675945A1C287DA30 CRC64;

  Query Match             25.8%;  Score 51;  DB 6;  Length 96;
  Best Local Similarity   41.7%;  Pred. No. 9;
  Matches   10;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      10 RDSRDPLWKGPAKLLWKGEGAVVI 33
           | |  | |:| |  ||   | :|:
Db      49 RRSSTPHWQGSAHPLWMSPGLLVL 72


RESULT 7
Q97KS3
ID   Q97KS3      PRELIMINARY;      PRT;    91 AA.
AC   Q97KS3;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Barstar-like protein ribonuclease (barnase) inhibitor.
GN   CAC0844.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007600; AAK78820.1; -.
DR   InterPro; IPR000468; Barstar.
DR   Pfam; PF01337; Barstar; 1.
DR   ProDom; PD029050; Barstar; 1.
KW   Complete proteome.
SQ   SEQUENCE   91 AA;  10821 MW;  AD4B022BC9FCCB8E CRC64;

  Query Match             25.5%;  Score 50.5;  DB 16;  Length 91;
  Best Local Similarity   37.9%;  Pred. No. 10;
  Matches   11;  Conservative    3;  Mismatches    8;  Indels    7;  Gaps    1;

Qy       5 FRVYYRDSRDPLW-------KGPAKLLWK 26
           |  ||  : | ||       | | |::||
Db      27 FPEYYGKNLDALWDCLTGETKMPLKIVWK 55


RESULT 8
Q71895
ID   Q71895      PRELIMINARY;      PRT;    44 AA.
AC   Q71895;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24447; AAA79545.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
```

```
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 9
Q71902
ID   Q71902       PRELIMINARY;      PRT;    44 AA.
AC   Q71902;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24448; AAA79552.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 10
Q71909
ID   Q71909       PRELIMINARY;      PRT;    44 AA.
AC   Q71909;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24449; AAA79559.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 11
Q71916
ID   Q71916       PRELIMINARY;      PRT;    44 AA.
AC   Q71916;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24450; AAA79566.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 12
Q71923
ID   Q71923       PRELIMINARY;      PRT;    44 AA.
AC   Q71923;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24451; AAA79573.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 13
Q71929
ID   Q71929      PRELIMINARY;      PRT;    44 AA.
AC   Q71929;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 1;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24452; AAA79579.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4875 MW;  01901EA27AAF3CDE CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 14
Q71936
ID   Q71936      PRELIMINARY;      PRT;    44 AA.
AC   Q71936;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 2;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24453; AAA79586.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4690 MW;  F0F71C7266DF3CC6 CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


RESULT 15
Q71942
ID   Q71942      PRELIMINARY;      PRT;    44 AA.
AC   Q71942;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pol protein (Fragment).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNTRL 2;
RX   MEDLINE=95287475; PubMed=7769682;
RA   Michael N.L., Chang G., d'Arcy L.A., Ehrenberg P.K., Mariani R.,
RA   Busch M.P., Birx D.L., Schwartz D.H.;
RT   "Defective accessory genes in a human immunodeficiency virus type 1-
RT   infected long-term survivor lacking recoverable virus.";
RL   J. Virol. 69:4228-4236(1995).
DR   EMBL; U24454; AAA79592.1; -.
DR   InterPro; IPR001037; Integrase_C.
DR   Pfam; PF00552; integrase; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   44 AA;  4762 MW;  F15C041266DF3CC6 CRC64;

  Query Match             25.3%;  Score 50;  DB 15;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      27 GEGAVVIQDN 36
           ||||||||||
Db       1 GEGAVVIQDN 10


Search completed: May  6, 2003, 14:58:17
Job time : 31 secs
```

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:55:39 ; Search time 11 Seconds
(without alignments)
135.741 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 12886

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38.5	19.4	96	1	YY02_SYNY3	P73836 synechocyst
2	38	19.2	83	1	VG03_BPMD2	O64199 mycobacteri
3	37.5	18.9	73	1	GBGB_HUMAN	P50152 homo sapien
4	36.5	18.4	97	1	IF1A_ARCFU	O29481 archaeoglob
5	36	18.2	36	1	YG50_HAEIN	P44281 haemophilus
6	36	18.2	73	1	GBG1_BOVIN	P02698 bos taurus
7	36	18.2	87	1	R35A_PYRAB	Q9v1p2 pyrococcus
8	36	18.2	87	1	R35A_PYRFU	Q8tzv6 pyrococcus
9	36	18.2	87	1	R35A_PYRWO	P20299 pyrococcus
10	36	18.2	89	1	YVAP_VACCC	P20525 vaccinia vi
11	35.5	17.9	89	1	BARS_BACAM	P11540 bacillus am
12	35.5	17.9	91	1	SR19_METTH	O26267 methanobact
13	35	17.7	76	1	A4_MACMU	P29216 macaca mula
14	35	17.7	83	1	VG7_SPVIR	P15898 spiroplasma
15	35	17.7	85	1	VG73_BPML5	Q05288 mycobacteri
16	35	17.7	87	1	A4_MACFA	P53601 macaca fasc
17	34.5	17.4	73	1	YB76_HAEIN	P44120 haemophilus
18	34.5	17.4	97	1	Y117_NPVOP	O10356 orgyia pseu
19	34	17.2	56	1	AMCI_APIME	P56682 apis mellif
20	34	17.2	65	1	LHA2_ECTHL	P80103 ectothiorho
21	34	17.2	66	1	YVED_VACCC	P20556 vaccinia vi
22	34	17.2	88	1	Y4OL_RHISN	P55597 rhizobium s
23	34	17.2	93	1	FPTB_PSEAE	P42513 pseudomonas
24	34	17.2	97	1	YBGE_ECOLI	P37343 escherichia
25	33.5	16.9	79	1	RS16_BUCAI	P57474 buchnera ap
26	33.5	16.9	82	1	RS16_YERPE	Q8zbu7 yersinia pe
27	33.5	16.9	92	1	RS19_CAUCR	Q9a8u9 caulobacter
28	33.5	16.9	93	1	RT19_MARPO	P26874 marchantia
29	33.5	16.9	94	1	ES65_MYCTU	O53942 mycobacteri
30	33.5	16.9	94	1	ES66_MYCTU	P95242 mycobacteri
31	33	16.7	65	1	PSAL_ANAVA	P31092 anabaena va
32	33	16.7	71	1	HXC5_NOTVI	P31262 notophthalm
33	33	16.7	75	1	MT1B_VICFA	Q41670 vicia faba
34	33	16.7	77	1	VG43_BPML5	Q05255 mycobacteri
35	33	16.7	83	1	Y67_BPT3	P20330 bacteriopha
36	33	16.7	92	1	RS19_RICCN	Q92gx0 rickettsia
37	33	16.7	92	1	RS19_RICPR	Q9zcq9 rickettsia
38	32.5	16.4	36	1	NUCM_SOLTU	P80264 solanum tub
39	32.5	16.4	92	1	RS19_AGRT5	Q8ue22 agrobacteri
40	32.5	16.4	94	1	IFA2_HALN1	Q9hp87 halobacteri
41	32.5	16.4	98	1	S10Z_HUMAN	Q8wxg8 homo sapien
42	32	16.2	17	1	PH4_PERAM	P82697 periplaneta
43	32	16.2	37	1	ES2B_RANES	P40846 rana escule
44	32	16.2	52	1	A1T3_HORSE	P38030 equus cabal
45	32	16.2	58	1	68MP_MOUSE	P56379 mus musculu

ALIGNMENTS

RESULT 1
YY02_SYNY3
ID YY02_SYNY3 STANDARD; PRT; 96 AA.
AC P73836;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ssr3402 precursor.
GN SSR3402.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; D90910; BAA17893.1; -.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 96 HYPOTHETICAL PROTEIN SSR3402.
SQ SEQUENCE 96 AA; 10673 MW; 74175E6D70068C3E CRC64;

Query Match 19.4%; Score 38.5; DB 1; Length 96;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```
Qy      10 RDSRDPL-WKGPAKLL 24
           : |  || | |||:||
Db      71 KPSPKPLGWAGPAELL 86
```

RESULT 2
VG03_BPMD2
ID VG03_BPMD2 STANDARD; PRT; 83 AA.
AC O64199;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 3 protein (GP3).

```
GN   3.
OS   Mycobacteriophage D29.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC   unclassified Siphoviridae.
OX   NCBI_TaxID=28369;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98300335; PubMed=9636706;
RA   Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT   "Genome structure of mycobacteriophage D29: implications for phage
RT   evolution.";
RL   J. Mol. Biol. 279:143-164(1998).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF022214; AAC18446.1; -.
SQ   SEQUENCE   83 AA;  8962 MW;  2B3D070B4C6F4E09 CRC64;

  Query Match             19.2%;  Score 38;  DB 1;  Length 83;
  Best Local Similarity   36.8%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      17 WKGPAKLLWKGEGAVVIQD 35
           |:|   | : |:| | : |
Db      42 WEGLEILEYSGDGQVEVSD 60


RESULT 3
GBGB_HUMAN
ID   GBGB_HUMAN      STANDARD;      PRT;    73 AA.
AC   P50152;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-11 subunit.
GN   GNG11 OR GNGT11.
OS   Homo sapiens (Human),
OS   Mus musculus (Mouse), and
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606, 10090, 10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND ISOPRENOID.
RC   SPECIES=Human; TISSUE=Testis;
RX   MEDLINE=95394940; PubMed=7665596;
RA   Ray K., Kunsch C., Bonner L.M., Robishaw J.D.;
RT   "Isolation of cDNA clones encoding eight different human G protein
RT   gamma subunits, including three novel forms designated the gamma 4,
RT   gamma 10, and gamma 11 subunits.";
RL   J. Biol. Chem. 270:21765-21771(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RA   Maggi L.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human; TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Kidney, and Tongue;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Rat; STRAIN=Sprague-Dawley;
RA   Costain W.J., Mishra R.K.;
RT   "Identification and cloning of rat G protein gamma 11 subunit.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE
CC       GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-
CC       EFFECTOR INTERACTION. INTERACTS WITH BETA-1 AND BETA-3, BUT NOT
CC       WITH BETA-2.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC   -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN ALL TISSUES TESTED
CC       EXCEPT FOR BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE G PROTEIN GAMMA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31384; AAC50206.1; -.
DR   EMBL; AC002076; AAB54051.1; -.
DR   EMBL; BC009709; AAH09709.1; -.
DR   EMBL; AK002765; BAB22340.1; -.
DR   EMBL; AK009529; BAB26342.1; -.
DR   EMBL; AF257110; AAF68984.1; -.
DR   HSSP; P02698; 1B9X.
DR   Genew; HGNC:4403; GNG11.
DR   MIM; 604390; -.
DR   InterPro; IPR001770; G-gamma.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00631; G-gamma; 1.
DR   PRINTS; PR00321; GPROTEING.
DR   ProDom; PD003783; G-gamma; 1.
DR   SMART; SM00224; GGL; 1.
DR   PROSITE; PS50058; G_PROTEIN_GAMMA; 1.
KW   Transducer; Prenylation; Lipoprotein; Multigene family.
FT   LIPID       70     70       FARNESYL.
FT   PROPEP      71     73       REMOVED IN MATURE FORM (BY SIMILARITY).
SQ   SEQUENCE   73 AA;  8481 MW;  2B13935E3AEFB9E8 CRC64;

  Query Match             18.9%;  Score 37.5;  DB 1;  Length 73;
  Best Local Similarity   37.1%;  Pred. No. 1.2e+02;
  Matches   13;  Conservative    7;  Mismatches   12;  Indels    3;  Gaps    2;

Qy       1 KIQNFRVYYRDSRDPLWKG--PAKLLWKGEGAVVI 33
             :|:|: :  |   ||| ||    |  :| :|: ||
Db      39 EIKNY-IEERSGEDPLVKGIPEDKNPFKEKGSCVI 72
```

```
RESULT 4
IF1A_ARCFU
ID   IF1A_ARCFU     STANDARD;      PRT;    97 AA.
AC   O29481;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable translation initiation factor 1A (aIF-1A).
GN   EIF1A OR AF0777.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: Seems to be required for maximal rate of protein
CC       biosynthesis. Enhances ribosome dissociation into subunits and
CC       stabilizes The binding of the initiator Met-tRNA(I) to 40 S
CC       ribosomal subunits (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE EIF-1A FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 S1-LIKE DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001051; AAB90469.1; ALT_INIT.
DR   TIGR; AF0777; -.
DR   InterPro; IPR001253; TIF_eIF-1A.
DR   Pfam; PF01176; eIF-1a; 1.
DR   ProDom; PD005579; TIF_eIF-1A; 1.
DR   TIGRFAMs; TIGR00523; eIF-1A; 1.
DR   PROSITE; PS01262; IF1A; 1.
DR   PROSITE; PS50832; S1_IF1_TYPE; 1.
KW   Initiation factor; Protein biosynthesis; Complete proteome.
FT   DOMAIN        8     82       S1-LIKE.
SQ   SEQUENCE   97 AA;  11385 MW;  12332D5E9CB82DA1 CRC64;

  Query Match             18.4%;  Score 36.5;  DB 1;  Length 97;
  Best Local Similarity   33.3%;  Pred. No. 2.2e+02;
  Matches   11;  Conservative    6;  Mismatches   13;  Indels    3;  Gaps    1;

Qy       4 NFRVYYRDSRDPLWKGPAKL---LWKGEGAVVI 33
           : :|    |  : | : | |:   :|  || |||
Db      32 HIKVRCEDGVERLARIPGKMRKKIWIREGDVVI 64


RESULT 5
YG50_HAEIN
ID   YG50_HAEIN     STANDARD;      PRT;    36 AA.
AC   P44281;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein HI1650.
GN   HI1650.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U32838; AAC23303.1; -.
DR   TIGR; HI1650; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   36 AA;  4285 MW;  248010DAD7898E33 CRC64;

  Query Match             18.2%;  Score 36;  DB 1;  Length 36;
  Best Local Similarity   42.9%;  Pred. No. 90;
  Matches    9;  Conservative    2;  Mismatches    8;  Indels    2;  Gaps    1;

Qy       2 IQNFRVYYRDSRDPLWKGPAK 22
           |||:   | : ||  | |  |
Db       7 IQNYHSTYNNIRD--WLGYQK 25


RESULT 6
GBG1_BOVIN
ID   GBG1_BOVIN     STANDARD;      PRT;    73 AA.
AC   P02698;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(T) gamma-T1 subunit (Transducin
DE   gamma chain).
GN   GNGT1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85063709; PubMed=6438626;
RA   Hurley J.B., Fong H.K.W., Teplow D.B., Dreyer W.J., Simon M.I.;
RT   "Isolation and characterization of a cDNA clone for the gamma subunit
RT   of bovine retinal transducin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:6948-6952(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85166247; PubMed=2984674;
RA   Yatsunami K., Pandya B.V., Oprian D.D., Khorana H.G.;
RT   "cDNA-derived amino acid sequence of the gamma subunit of GTPase from
```

```
RT   bovine rod outer segments.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1936-1940(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93272877; PubMed=8500562;
RA   Tao L., Pandey S., Simon M.I., Fong H.K.;
RT   "Structure of the bovine transducin gamma subunit gene and analysis
RT   of promoter function in transgenic mice.";
RL   Exp. Eye Res. 56:497-507(1993).
RN   [4]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE=85046503; PubMed=6149748;
RA   van Dop C., Medynski D.C., Sullivan K., Wu A.M., Fung B.K.-K.,
RA   Bourne H.R.;
RT   "Partial cDNA sequence of the gamma subunit of transducin.";
RL   Biochem. Biophys. Res. Commun. 124:250-255(1984).
RN   [5]
RP   SEQUENCE OF 1-69.
RX   MEDLINE=85076983; PubMed=3917402;
RA   Ovchinnikov Y.A., Lipkin V.M., Shuvaeva T.M., Bogachuk A.P.,
RA   Shemyakin V.V.;
RT   "Complete amino acid sequence of gamma-subunit of the GTP-binding
RT   protein from cattle retina.";
RL   FEBS Lett. 179:107-110(1985).
RN   [6]
RP   ISOPRENOID.
RX   MEDLINE=90348966; PubMed=2385292;
RA   Fukuda Y., Takao T., Ohguro H., Yoshizawa T., Akino T.,
RA   Shimonishi Y.;
RT   "Farnesylated gamma-subunit of photoreceptor G protein indispensable
RT   for GTP-binding.";
RL   Nature 346:658-660(1990).
RN   [7]
RP   ISOPRENOID.
RX   MEDLINE=91236727; PubMed=1903391;
RA   Sanford J., Codina J., Birnbaumer L.;
RT   "Gamma-subunits of G proteins, but not their alpha- or beta-subunits,
RT   are polyisoprenylated. Studies on post-translational modifications
RT   using in vitro translation with rabbit reticulocyte lysates.";
RL   J. Biol. Chem. 266:9570-9579(1991).
RN   [8]
RP   ISOPRENOID.
RX   MEDLINE=91017567; PubMed=2217200;
RA   Lai R.K., Perez-Sala D., Canada F.J., Rando R.R.;
RT   "The gamma subunit of transducin is farnesylated.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7673-7677(1990).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF BETA-GAMMA DIMER.
RX   MEDLINE=96149254; PubMed=8552196;
RA   Sondek J., Bohm A., Lambright D.G., Hamm H.E., Sigler P.B.;
RT   "Crystal structure of a G-protein beta gamma dimer at 2.1-A
RT   resolution.";
RL   Nature 379:369-374(1996).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH PHOSDUCIN.
RX   MEDLINE=98416696; PubMed=9739091;
RA   Loew A., Ho Y.K., Blundell T., Bax B.;
RT   "Phosducin induces a structural change in transducin beta gamma.";
RL   Structure 6:1007-1019(1998).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE
CC       GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-
CC       EFFECTOR INTERACTION.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC   -!- TISSUE SPECIFICITY: RETINAL ROD OUTER SEGMENT.
CC   -!- SIMILARITY: BELONGS TO THE G PROTEIN GAMMA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03255; AAA30794.1; -.
DR   EMBL; K02199; AAA30793.1; -.
DR   EMBL; S62031; AAB26895.1; -.
DR   EMBL; S62029; AAB26895.1; JOINED.
DR   EMBL; K02436; AAA30788.1; -.
DR   PIR; A03153; RGBOGT.
DR   PDB; 1TBG; 01-APR-97.
DR   PDB; 1A0R; 16-FEB-99.
DR   PDB; 1B9X; 23-FEB-99.
DR   PDB; 1B9Y; 23-FEB-99.
DR   InterPro; IPR001770; G-gamma.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00631; G-gamma; 1.
DR   PRINTS; PR00321; GPROTEING.
DR   ProDom; PD003783; G-gamma; 1.
DR   SMART; SM00224; GGL; 1.
DR   PROSITE; PS50058; G_PROTEIN_GAMMA; 1.
KW   Transducer; Prenylation; Lipoprotein; Multigene family; 3D-structure.
FT   INIT_MET      0      0
FT   DISULFID     35     36
FT   LIPID        70     70       FARNESYL.
FT   PROPEP       71     73       REMOVED IN MATURE FORM.
SQ   SEQUENCE   73 AA;  8413 MW;  B1743E36F4BD2505 CRC64;

  Query Match             18.2%;  Score 36;  DB 1;  Length 73;
  Best Local Similarity   42.1%;  Pred. No. 1.9e+02;
  Matches   16;  Conservative    2;  Mismatches   10;  Indels   10;  Gaps    4;

Qy       3 QNFRVYY--RDSRDPLWKG------PAKLLWKGEGAVV 32
           : || |    |   ||| ||      | | | || | |:
Db      37 EEFRDYVEERSGEDPLVKGIPEDKNPFKEL-KG-GCVI 72


RESULT 7
R35A_PYRAB
ID   R35A_PYRAB     STANDARD;      PRT;    87 AA.
AC   Q9V1P2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L35Ae.
GN   RPL35AE OR PAB7092.
OS   Pyrococcus abyssi.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=29292;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GE5 / Orsay;
RA   Heilig R.;
RT   "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT   structure and evolution.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE L35AE FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ248284; CAB49307.1; -.
DR   InterPro; IPR001780; Ribosomal_L35AE.
DR   Pfam; PF01247; Ribosomal_L35Ae; 1.
DR   ProDom; PD012670; Ribosomal_L35AE; 1.
DR   PROSITE; PS01105; RIBOSOMAL_L35AE; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   87 AA;  9736 MW;  79A625A3AF744958 CRC64;
```

```
  Query Match             18.2%;  Score 36;  DB 1;  Length 87;
  Best Local Similarity   46.2%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      16 LWKGPAKLLWKGE 28
           :|| |:  | ||:
Db      44 IWKSPSGKLLKGK 56


RESULT 8
R35A_PYRFU
ID   R35A_PYRFU     STANDARD;      PRT;    87 AA.
AC   Q8TZV6;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L35Ae.
GN   RPL35AE OR PF1872.
OS   Pyrococcus furiosus.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=2261;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA   Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT   "The complete sequence of the Pyrococcus furiosus genome.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE L35AE FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE010282; AAL81996.1; -.
DR   InterPro; IPR001780; Ribosomal_L35AE.
DR   Pfam; PF01247; Ribosomal_L35Ae; 1.
DR   ProDom; PD012670; Ribosomal_L35AE; 1.
DR   PROSITE; PS01105; RIBOSOMAL_L35AE; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   87 AA;  9735 MW;  F5C425AAA3A09247 CRC64;

  Query Match             18.2%;  Score 36;  DB 1;  Length 87;
  Best Local Similarity   46.2%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      16 LWKGPAKLLWKGE 28
           ||| |:  : ||:
Db      44 LWKSPSGKILKGK 56


RESULT 9
R35A_PYRWO
ID   R35A_PYRWO     STANDARD;      PRT;    87 AA.
AC   P20299;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   50S ribosomal protein L35Ae.
GN   RPL35AE.
OS   Pyrococcus woesei.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=2262;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 3773;
RX   MEDLINE=90330536; PubMed=2165475;
RA   Zwickl P., Fabry S., Bogedain C., Haas A., Hensel R.;
RT   "Glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic
RT   archaebacterium Pyrococcus woesei: characterization of the enzyme,
RT   cloning and sequencing of the gene, and expression in Escherichia
RT   coli.";
RL   J. Bacteriol. 172:4329-4338(1990).
CC   -!- SIMILARITY: BELONGS TO THE L35AE FAMILY OF RIBOSOMAL PROTEINS.
DR   PIR; S10651; QQQYYW.
DR   InterPro; IPR001780; Ribosomal_L35AE.
DR   Pfam; PF01247; Ribosomal_L35Ae; 1.
DR   ProDom; PD012670; Ribosomal_L35AE; 1.
DR   PROSITE; PS01105; RIBOSOMAL_L35AE; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   87 AA;  9666 MW;  B5D524BBA3A148EC CRC64;

  Query Match             18.2%;  Score 36;  DB 1;  Length 87;
  Best Local Similarity   46.2%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      16 LWKGPAKLLWKGE 28
           ||| |:  : ||:
Db      44 LWKSPSGKILKGK 56


RESULT 10
YVAP_VACCC
ID   YVAP_VACCC     STANDARD;      PRT;    89 AA.
AC   P20525;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 9.9 kDa protein.
GN   A ORF P.
OS   Vaccinia virus (strain Copenhagen).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10249;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91021027; PubMed=2219722;
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "The complete DNA sequence of vaccinia virus.";
RL   Virology 179:247-266(1990).
RN   [2]
RP   COMPLETE GENOME.
RA   Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA   Paoletti E.;
RT   "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL   Virology 179:517-563(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35027; AAA48170.1; -.
DR   PIR; C42525; C42525.
KW   Hypothetical protein.
SQ   SEQUENCE   89 AA;  9909 MW;  399EA9270DF3E75A CRC64;

  Query Match             18.2%;  Score 36;  DB 1;  Length 89;
  Best Local Similarity   35.7%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      19 GPAKLLWKGEGAVV 32
           ||:|: |  :| ::
Db      24 GPSKIGWLDKGFII 37
```

```
RESULT 11
BARS_BACAM
ID   BARS_BACAM      STANDARD;      PRT;    89 AA.
AC   P11540;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Barstar (Ribonuclease inhibitor).
OS   Bacillus amyloliquefaciens.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1390;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89012012; PubMed=3050134;
RA   Hartley R.W.;
RT   "Barnase and barstar. Expression of its cloned inhibitor permits
RT   expression of a cloned ribonuclease.";
RL   J. Mol. Biol. 202:913-915(1988).
RN   [2]
RP   REVIEW.
RX   MEDLINE=90162921; PubMed=2696173;
RA   Hartley R.W.;
RT   "Barnase and barstar: two small proteins to fold and fit together.";
RL   Trends Biochem. Sci. 14:450-454(1989).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH BARNASE.
RA   Guillet V., Lapthorn A., Hartley R.W., Mauguen Y.;
RT   "Recognition between a bacterial ribonuclease, barnase, and its
RT   natural inhibitor, barstar.";
RL   Structure 1:165-177(1993).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF COMPLEX WITH RNASE SA.
RX   MEDLINE=98437624; PubMed=9757110;
RA   Sevcik J., Urbanikova L., Dauter Z., Wilson K.S.;
RT   "Recognition of RNase Sa by the inhibitor barstar: structure of the
RT   complex at 1.7 A resolution.";
RL   Acta Crystallogr. D 54:954-963(1998).
RN   [5]
RP   STRUCTURE BY NMR.
RX   MEDLINE=94009694; PubMed=8405454;
RA   Lubienski M.J., Bycroft M., Jones D.N.M., Fersht A.R.;
RT   "Assignment of the backbone 1H and 15N NMR resonances and secondary
RT   structure characterization of barstar.";
RL   FEBS Lett. 332:81-87(1993).
RN   [6]
RP   STRUCTURE BY NMR.
RX   MEDLINE=94318630; PubMed=8043574;
RA   Lubienski M.J., Bycroft M., Freund S.M.V., Fersht A.R.;
RT   "Three-dimensional solution structure and 13C assignments of barstar
RT   using nuclear magnetic resonance spectroscopy.";
RL   Biochemistry 33:8866-8877(1994).
CC   -!- FUNCTION: INHIBITOR OF THE RIBONUCLEASE BARNASE. FORMS A ONE-TO-
CC       ONE NON-COVALENT COMPLEX.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15545; CAA33551.1; -.
DR   PIR; S01373; S01373.
DR   PDB; 1BRS; 31-JUL-94.
DR   PDB; 1BTA; 31-JUL-94.
DR   PDB; 1BTB; 31-JUL-94.
DR   PDB; 1AB7; 04-SEP-97.
DR   PDB; 1A19; 08-APR-98.
DR   PDB; 1B27; 09-DEC-98.
DR   PDB; 1B2S; 09-DEC-98.
DR   PDB; 1B2U; 09-DEC-98.
DR   PDB; 1B3S; 09-DEC-98.
DR   PDB; 1AY7; 02-MAR-99.
DR   InterPro; IPR000468; Barstar.
DR   Pfam; PF01337; Barstar; 1.
DR   ProDom; PD029050; Barstar; 1.
KW   3D-structure.
FT   INIT_MET       0       0
SQ   SEQUENCE   89 AA;  10212 MW;  3AC7E76A9C43A505 CRC64;

  Query Match             17.9%;  Score 35.5;  DB 1;  Length 89;
  Best Local Similarity   30.8%;  Pred. No. 2.8e+02;
  Matches    8;  Conservative    4;  Mismatches    7;  Indels    7;  Gaps    1;

Qy       8 YYRDSRDPLW-------KGPAKLLWK 26
           || :: | ||        : |  | |:
Db      29 YYGENLDALWDCLTGWVEYPLVLEWR 54


RESULT 12
SR19_METTH
ID   SR19_METTH      STANDARD;      PRT;    91 AA.
AC   O26267;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Signal recognition particle 19 kDa protein (SRP19).
GN   SRP19 OR MTH165.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -!- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY, BINDS DIRECTLY TO
CC       7S RNA AND MEDIATES BINDING OF THE 54 KDA SUBUNIT OF THE SRP (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: ARCHAEAL SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA
CC       MOLECULE OF 300 NUCLEOTIDES AND TWO PROTEIN SUBUNITS: SRP54 AND
CC       SRP19 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000804; AAB84671.1; -.
DR   InterPro; IPR002778; SRP19.
DR   Pfam; PF01922; SRP19; 1.
DR   ProDom; PD006609; SRP19; 1.
KW   Signal recognition particle; RNA-binding; Ribonucleoprotein;
KW   Complete proteome.
SQ   SEQUENCE   91 AA;  10535 MW;  BFD64D1A20B48141 CRC64;

  Query Match             17.9%;  Score 35.5;  DB 1;  Length 91;
  Best Local Similarity   31.8%;  Pred. No. 2.9e+02;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    7;  Gaps    1;
```

```
Qy      15 PLWKGPAKLLWKGEGAVVIQDN 36
           | |       |:  | ||:: |
Db      57 PSW-------WESSGRVVVEYN 71
```

```
RESULT 13
A4_MACMU
ID   A4_MACMU       STANDARD;      PRT;    76 AA.
AC   P29216;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Alzheimer's disease amyloid A4 protein (Fragment).
GN   APP.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Koo E.H., Sisodia S.S., Price D.L.;
RL   Submitted (JUL-1989) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC       INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC       G(O) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
CC       APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
CC       SPLICING (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15985; CAA34116.1; -.
DR   PIR; S06678; S06678.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR001868; A4_APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; PARTIAL.
DR   PROSITE; PS00320; A4_INTRA; PARTIAL.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Glycoprotein; Amyloid; Neurone; Alternative splicing;
KW   Serine protease inhibitor.
FT   NON_TER       1      1
FT   DOMAIN        1     76       BPTI/KUNITZ INHIBITOR.
FT   ACT_SITE     13     14       REACTIVE BOND.
FT   DISULFID      3     53       BY SIMILARITY.
FT   DISULFID     12     36       BY SIMILARITY.
FT   DISULFID     28     49       BY SIMILARITY.
FT   NON_TER      76     76
SQ   SEQUENCE   76 AA;  8527 MW;  492BF3069AB082A1 CRC64;

  Query Match           17.7%;  Score 35;  DB 1;  Length 76;
  Best Local Similarity  42.9%;  Pred. No. 2.8e+02;
  Matches    6;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      10 RDSRDPLWKGPAKL 23
           : :|:|| : | ||
Db      63 KTTREPLTRDPVKL 76
```

```
RESULT 14
VG7_SPV1R
ID   VG7_SPV1R      STANDARD;      PRT;    83 AA.
AC   P15898;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Gene 7 protein.
GN   7.
OS   Spiroplasma virus SpV1-R8A2 B.
OC   Viruses; ssDNA viruses; Inoviridae; Plectrovirus.
OX   NCBI_TaxID=10854;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90206799; PubMed=2320423;
RA   Renaudin J., Aullo P., Vignault J.C., Bove J.M.;
RT   "Complete nucleotide sequence of the genome of Spiroplasma citri
RT   virus SpV1-R8A2 B.";
RL   Nucleic Acids Res. 18:1293-1293(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51344; CAA35731.1; -.
SQ   SEQUENCE   83 AA;  9326 MW;  7F8286216CF74A9A CRC64;

  Query Match           17.7%;  Score 35;  DB 1;  Length 83;
  Best Local Similarity  45.5%;  Pred. No. 3.1e+02;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      16 LWKGPAKLLWK 26
           :| | :  |||
Db      28 IWTGLSSALWK 38
```

```
RESULT 15
VG73_BPML5
ID   VG73_BPML5     STANDARD;      PRT;    85 AA.
AC   Q05288;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Gene 73 protein (GP73).
GN   73.
OS   Mycobacteriophage L5.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC   L5-like viruses.
OX   NCBI_TaxID=31757;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93211282; PubMed=8459766;
RA   Hatfull G.F., Sarkis G.J.;
RT   "DNA sequence, structure and gene expression of mycobacteriophage L5:
RT   a phage system for mycobacterial genetics.";
RL   Mol. Microbiol. 7:395-405(1993).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z18946; CAA79449.1; -.
DR   PIR; S31018; S31018.
```

SQ SEQUENCE 85 AA; 10250 MW; 5032B5A1400FF4A7 CRC64;

Query Match 17.7%; Score 35; DB 1; Length 85;
Best Local Similarity 32.4%; Pred. No. 3.2e+02;
Matches 11; Conservative 4; Mismatches 13; Indels 6; Gaps 2;

```
Qy       2 IQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQD 35
           |:: | ::|   |  ||       |  | | : ||
Db      28 IKDLRKWFRADHDQ-WK-----TWHDEPAPFEQD 55
```

Search completed: May 6, 2003, 14:57:41
Job time : 12 secs

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:56:24 ; Search time 16 Seconds
(without alignments)
216.302 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 37666

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	53.5	27.0	70	2	A69842	hypothetical prote
2	50.5	25.5	91	2	A97004	barstar-like prote
3	50	25.3	44	2	T09381	pol polyprotein -
4	48.5	24.5	95	2	A81176	ribonuclease inhib
5	45	22.7	82	2	B95846	hypothetical prote
6	42	21.2	76	2	AD2538	hypothetical prote
7	41.5	21.0	96	2	G83600	hypothetical prote
8	41	20.7	94	2	A69848	hypothetical prote
9	40.5	20.5	68	2	AG3217	hypothetical prote
10	39.5	19.9	63	2	S26796	Ig heavy chain V r
11	39.5	19.9	87	2	T03193	hypothetical prote
12	39.5	19.9	88	2	T36458	hypothetical prote
13	39.5	19.9	96	2	T36605	hypothetical prote
14	39	19.7	62	2	AI2304	hypothetical prote
15	39	19.7	64	2	T05933	probable 3-methyl-
16	39	19.7	79	2	T37100	hypothetical prote
17	39	19.7	96	2	C64900	outer membrane por
18	38.5	19.4	61	2	C75321	preprotein translo
19	38.5	19.4	62	2	T26847	hypothetical prote
20	38.5	19.4	66	2	S32027	Sp12 protein homol
21	38.5	19.4	96	2	A71151	hypothetical prote
22	38.5	19.4	96	2	S75031	hypothetical prote
23	38	19.2	60	2	H81170	conserved hypothet
24	38	19.2	81	2	B95286	hypothetical prote
25	38	19.2	83	2	C72800	gp3 protein - Myco
26	38	19.2	96	2	H81147	hypothetical prote
27	38	19.2	100	2	F95369	hypothetical prote
28	37.5	18.9	60	2	H97179	hypothetical prote
29	37.5	18.9	67	2	AF1915	CAB/ELIP/HLIP supe
30	37.5	18.9	73	2	I39159	GTP-binding regula
31	37.5	18.9	88	2	T17532	hypothetical prote
32	37	18.7	50	2	S68843	sucrose 1F-fructos
33	37	18.7	56	2	B69185	hypothetical prote
34	37	18.7	76	2	B97816	hypothetical prote
35	37	18.7	77	2	A83468	hypothetical prote
36	37	18.7	82	2	H84046	hypothetical prote
37	37	18.7	94	2	H84041	hypothetical prote
38	36.5	18.4	62	2	D69045	hypothetical prote
39	36.5	18.4	88	2	A69347	translation initia
40	36.5	18.4	97	2	E90278	hypothetical prote
41	36.5	18.4	98	2	F90314	hypothetical prote
42	36.5	18.4	99	2	E90338	hypothetical prote
43	36	18.2	36	2	C64039	hypothetical prote
44	36	18.2	46	2	F87527	hypothetical prote
45	36	18.2	52	2	S21185	beta-fructofuranos

ALIGNMENTS

RESULT 1
A69842
hypothetical protein yitX - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: A69842
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69842
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-70 <KUN>
A;Cross-references: GB:Z99109; GB:AL009126; NID:g2633260; PIDN:CAB12957.1; PID:e1183119
A;Experimental source: strain 168
C;Genetics:
A;Gene: yitX

Query Match 27.0%; Score 53.5; DB 2; Length 70;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

```
Qy       6 RVYYRDSRDPLWKGPAKLLWKGEGAVVI 33
           |: ||  | |  |   |::|| | ||::
Db      44 RLLYR-RRPPCRKTGKKIIWKNEDAVMV 70
```

RESULT 2
A97004
barstar-like protein ribonuclease (barnase) inhibitor [imported] - Clostridium acetobut
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: A97004
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: A97004

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK78820.1; PID:g15023737; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0844

Query Match 25.5%; Score 50.5; DB 2; Length 91;
Best Local Similarity 37.9%; Pred. No. 6.2;
Matches 11; Conservative 3; Mismatches 8; Indels 7; Gaps 1;

Qy 5 FRVYYRDSRDPLW-------KGPAKLLWK 26
Db 27 FPEYYGKNLDALWDCLTGETKMPLKIVWK 55

RESULT 3
T09381
pol polyprotein - human immunodeficiency virus type 1 (isolate cntrl 1) (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate cntrl 1
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
C;Accession: T09381
R;Michael, N.L.; Chang, G.; d'ARCY, L.A.; Ehrenberg, P.K.; Mariani, R.; Busch, M.P.; Bir
J. Virol. 69, 4228-4236, 1995
A;Title: Defective accessory genes in a human immunodeficiency virus type 1-infected lon
A;Reference number: Z16654; MUID:95287475; PMID:7769682
A;Accession: T09381
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-44 <MIC>
A;Cross-references: EMBL:U24451; NID:g829440; PID:g829441
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; immunodeficiency; polyprotein

Query Match 25.3%; Score 50; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 GEGAVVIQDN 36
Db 1 GEGAVVIQDN 10

RESULT 4
A81176
ribonuclease inhibitor barstar NMB0646 [imported] - Neisseria meningitidis (strain MC58
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: A81176
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81176
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <TET>
A;Cross-references: GB:AE002419; GB:AE002098; NID:g7225863; PIDN:AAF41067.1; PID:g722587
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0646

Query Match 24.5%; Score 48.5; DB 2; Length 95;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 11; Conservative 6; Mismatches 9; Indels 7; Gaps 1;

Qy 1 KIQNFRVYYRDSRDPLW-------KGPAKLLWK 26
Db 22 KIFSIQDYYGNNLDALWDLLSTNVERPITLVWK 54

RESULT 5
B95846
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSym
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: B95846
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing en
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: B95846
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-82 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC48434.1; PID:g15139906; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20031
A;Genome: plasmid

Query Match 22.7%; Score 45; DB 2; Length 82;
Best Local Similarity 37.5%; Pred. No. 32;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 12 SRDPLWKGPAKLLWKGEGAVVIQD 35
Db 20 SKDGHWRDPRPRLAAADGQIVITD 43

RESULT 6
AD2538
hypothetical protein asl7591 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120b
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AD2538
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-76 <KUR>
A;Cross-references: GB:AP003602; PIDN:BAB77234.1; PID:g17134676; GSPDB:GN00181
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl7591
A;Genome: plasmid

Query Match 21.2%; Score 42; DB 2; Length 76;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 9 YRDSRDPLWKGPAKL 23
Db 3 HQSSQDSWWRSPAKI 17

RESULT 7
G83600
hypothetical protein PA0369 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: G83600
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83600
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-96 <STO>
A;Cross-references: GB:AE004474; GB:AE004091; NID:g9946210; PIDN:AAG03758.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0369

Query Match 21.0%; Score 41.5; DB 2; Length 96;
Best Local Similarity 32.1%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

```
Qy      14 DRLWKGPAKL-----LWKGEGAVVIQDN 36
           | ||:  | |      :| | |  :: |:  .
Db      35 DDLWQMAASLACFGEVWSGAGEDILLDD 62
```

RESULT 8
A69848
hypothetical protein yjcQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: A69848
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69848
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-94 <KUN>
A;Cross-references: GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13052.1; PID:e1183215;
A;Experimental source: strain 168
C;Genetics:
A;Gene: yjcQ

Query Match 20.7%; Score 41; DB 2; Length 94;
Best Local Similarity 45.5%; Pred. No. 1.4e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```
Qy       7 VYYRDSRDPLWKGPAKLLWKGE 28
           |:| | |  |:|   :|  |||
Db      49 VHYSDDRPHLYKLGPELTEKGE 70
```

RESULT 9
AG3217
hypothetical protein Atu5470 [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AG3217
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AG3217
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-68 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL46157.1; PID:g17743927; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5470
A;Genome: plasmid

Query Match 20.5%; Score 40.5; DB 2; Length 68;
Best Local Similarity 37.0%; Pred. No. 1.1e+02;
Matches 10; Conservative 4; Mismatches 6; Indels 7; Gaps 2;

```
Qy       9 YRDSRD-PLW------KGPAKLLWKGE 28
             ::| |  |||      :| |:  | ||
Db      20 WKDIRTAPLWRINGTGRGDAQSAWAGE 46
```

RESULT 10
S26796
Ig heavy chain V region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C;Accession: S26796
R;Mortari, F.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
Eur. J. Immunol. 22, 241-245, 1992
A;Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene fam
A;Reference number: S26786; MUID:92111632; PMID:1730251
A;Accession: S26796
A;Molecule type: mRNA
A;Residues: 1-63 <MOR>
A;Cross-references: EMBL:X61018; NID:g32785; PIDN:CAA43352.1; PID:g1335121
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 19.9%; Score 39.5; DB 2; Length 63;
Best Local Similarity 31.2%; Pred. No. 1.4e+02;
Matches 10; Conservative 9; Mismatches 8; Indels 5; Gaps 3;

```
Qy       1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVV 32
             ::::  |||  ::|| |:       | |:| :|
Db      33 RVEDTAVYY-CAKDPNWEDAP---W-GQGTLV 59
```

RESULT 11
T03193
hypothetical protein 87 - rice mitochondrion
C;Species: mitochondrion Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T03193
R;Itadani, H.; Wakasugi, T.; Sugita, M.; Sugiura, M.; Nakazono, M.; Hirai, A.
Plant Cell Physiol. 35, 1239-1244, 1994
A;Title: Nucleotide sequence of a 28-kbp portion of rice mitochondrial DNA: the existen
A;Reference number: Z14841; MUID:95211382; PMID:7545979
A;Accession: T03193
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-87 <ITA>
A;Cross-references: EMBL:D32052; NID:g769704; PIDN:BAA06814.1; PID:g769710
A;Experimental source: cultivar Nipponbare

```
C;Genetics:
A;Genome: mitochondrion
C;Keywords: mitochondrion

  Query Match             19.9%;  Score 39.5;  DB 2;  Length 87;
  Best Local Similarity   31.0%;  Pred. No. 2e+02;
  Matches   9;  Conservative    7;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       3 QNFRVYYRDSRDPLWKGPAKLLWKGEGAV 31
           ||  : ||: |    || ::::|   | :
Db      31 QNILI-YRNQRPRHCKGGSRMIWIRTGCL 58


RESULT 12
T36458
hypothetical protein SCF43A.35c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36458
R;Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21598
A;Accession: T36458
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-88 <SEE>
A;Cross-references: EMBL:AL096837; PIDN:CAB48922.1; GSPDB:GN00070; SCOEDB:SCF43A.35c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCF43A.35c

  Query Match             19.9%;  Score 39.5;  DB 2;  Length 88;
  Best Local Similarity   42.9%;  Pred. No. 2e+02;
  Matches    9;  Conservative    3;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      10 RDSRDPLWKGPAKLLWKGEGA 30
           || |: | ||  |: |   |:
Db      18 RDDRERLKKGD-KVTWSSHGS 37


RESULT 13
T36605
hypothetical protein SCH24.40c - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36605
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21575
A;Accession: T36605
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-96 <OLI>
A;Cross-references: EMBL:AL049826; PIDN:CAB42746.1; GSPDB:GN00070; SCOEDB:SCH24.40c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH24.40c

  Query Match             19.9%;  Score 39.5;  DB 2;  Length 96;
  Best Local Similarity   47.6%;  Pred. No. 2.3e+02;
  Matches   10;  Conservative    4;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       1 KIQNFRVYYRDSRDP-LWKGP 20
           |:|:| | : :| || |   ||
Db      25 KVQHFFVCHLESMDPSLRHGP 45


RESULT 14
AI2304
hypothetical protein asr3992 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AI2304
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2304
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-62 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB75691.1; PID:g17133127; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asr3992
C;Superfamily: conserved hypothetical protein ycf9

  Query Match             19.7%;  Score 39;  DB 2;  Length 62;
  Best Local Similarity   53.8%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      17 WKGPAKLLWKGEG 29
           |    |||| | |
Db      33 WVESKKLLWLGSG 45


RESULT 15
T05933
probable 3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) alpha chain - b
N;Alternate names: branched-chain alpha-keto acid decarboxylase complex E1-alpha chain
C;Species: Hordeum vulgare (barley)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C;Accession: T05933
R;Hess, W.R.; Golz, R.R.; Boerner, T.
Plant Sci. 133, 191-201, 1998
A;Title: Analysis of randomly selcelted cDNAs reveals the expression of stress- and def
A;Reference number: Z15411
A;Accession: T05933
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-64 <HES>
A;Cross-references: EMBL:AJ222787; PIDN:CAA10992.1
A;Experimental source: cv. Haisa, leaf
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bi
C;Keywords: mitochondrion; oxidoreductase

  Query Match             19.7%;  Score 39;  DB 2;  Length 64;
  Best Local Similarity   46.2%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      15 PLWKGPAKLLWKG 27
           | :: |  |||:|
Db      11 PQYREPGVLLWRG 23


Search completed: May  6, 2003, 14:58:40
Job time : 18 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:29 ; Search time 16 Seconds
(without alignments)
194.145 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 125019

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	54	27.3	10	10	US-09-894-018-161	Sequence 161, App
2	54	27.3	15	10	US-09-894-018-192	Sequence 192, App
3	49	24.7	9	9	US-10-106-487-36	Sequence 36, Appl
4	49	24.7	14	9	US-10-062-710-23	Sequence 23, Appl
5	46.5	23.5	96	9	US-09-764-868-871	Sequence 871, App
6	46	23.2	38	9	US-10-082-830-200	Sequence 200, App
7	45.5	23.0	79	9	US-10-125-540-534	Sequence 534, App
8	45.5	23.0	79	10	US-09-764-870-534	Sequence 534, App
9	45	22.7	99	9	US-09-764-891-4912	Sequence 4912, Ap
10	44.5	22.5	93	9	US-09-764-891-5337	Sequence 5337, Ap
11	43	21.7	75	10	US-09-764-877-1243	Sequence 1243, Ap
12	43	21.7	86	9	US-09-989-920-243	Sequence 243, App
13	41	20.7	38	9	US-10-011-585A-230	Sequence 230, App
14	41	20.7	53	10	US-09-864-761-36417	Sequence 36417, A
15	41	20.7	94	10	US-09-734-569-120	Sequence 120, App
16	40.5	20.5	77	10	US-09-864-761-48829	Sequence 48829, A
17	40.5	20.5	86	9	US-09-796-692-2169	Sequence 2169, Ap
18	40.5	20.5	86	9	US-10-040-862-2169	Sequence 2169, Ap
19	40	20.2	31	10	US-09-822-540A-2	Sequence 2, Appli
20	40	20.2	31	10	US-09-938-700-2	Sequence 2, Appli
21	40	20.2	34	10	US-09-938-700-6	Sequence 6, Appli
22	40	20.2	38	9	US-10-185-050-34	Sequence 34, Appl
23	40	20.2	39	10	US-09-925-297-756	Sequence 756, App
24	40	20.2	60	10	US-09-864-761-49088	Sequence 49088, A
25	40	20.2	63	10	US-09-822-540A-1	Sequence 1, Appli
26	40	20.2	74	10	US-09-925-300-1663	Sequence 1663, Ap
27	40	20.2	91	9	US-10-001-857-198	Sequence 198, App
28	40	20.2	93	9	US-10-091-504-760	Sequence 760, App
29	40	20.2	93	10	US-09-764-869-760	Sequence 760, App
30	39.5	19.9	57	10	US-09-879-957-213	Sequence 213, App
31	39.5	19.9	63	9	US-09-925-664-34	Sequence 34, Appl
32	39.5	19.9	70	9	US-10-007-280A-210	Sequence 210, App
33	39	19.7	15	10	US-09-894-018-207	Sequence 207, App
34	39	19.7	49	10	US-09-864-761-37881	Sequence 37881, A
35	39	19.7	51	9	US-10-001-835-202	Sequence 202, App
36	39	19.7	66	9	US-09-796-692-2165	Sequence 2165, Ap
37	39	19.7	66	9	US-10-040-862-2165	Sequence 2165, Ap
38	39	19.7	67	10	US-09-864-761-33604	Sequence 33604, A
39	39	19.7	81	10	US-09-864-761-47481	Sequence 47481, A
40	39	19.7	88	10	US-09-864-761-46269	Sequence 46269, A
41	38.5	19.4	38	9	US-10-185-050-18	Sequence 18, Appl
42	38.5	19.4	50	10	US-09-864-761-38860	Sequence 38860, A
43	38	19.2	47	9	US-09-809-391-435	Sequence 435, App
44	38	19.2	56	9	US-09-764-891-3292	Sequence 3292, Ap
45	38	19.2	56	10	US-09-908-711-116	Sequence 116, App

ALIGNMENTS

RESULT 1
US-09-894-018-161
; Sequence 161, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Transgenic mouse
US-09-894-018-161

Query Match 27.3%; Score 54; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYR 10
||||||||||
Db 1 KIQNFRVYYR 10

RESULT 2
US-09-894-018-192

```
; Sequence 192, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
;  APPLICANT: EPIMMUNE, Inc.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Chestnut, Robert
;  APPLICANT:  Livingston, Brian
;  APPLICANT:  Baker, Denisw
;  APPLICANT:  Newman, Mark
;  APPLICANT:  Brown, David
;  TITLE OF INVENTION: METODS AND SYSTEM FOR OPTIMIZING
;  TITLE OF INVENTION:  MINIGENES AND PEPTIDES THEREBY
;  FILE REFERENCE: 39963-20033.00
;  CURRENT APPLICATION NUMBER: US/09/894,018
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: PCT/US00/35568
;  PRIOR FILING DATE: 2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/173,390
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: US 60/284,221
;  PRIOR FILING DATE: 2001-04-16
;  NUMBER OF SEQ ID NOS: 368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 192
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Transgenic mouse
US-09-894-018-192

  Query Match             27.3%;  Score 54;  DB 10;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYR 10
           ||||||||||
Db       6 KIQNFRVYYR 15


RESULT 3
US-10-106-487-36
; Sequence 36, Application US/10106487
; Patent No. US20020164721A1
; GENERAL INFORMATION:
;  APPLICANT: FIRAT, HUSEYIN
;  APPLICANT:  LEMONNIER, FRANCOIS
;  APPLICANT:  LANGLADE-DEMOYEN, PIERRE
;  APPLICANT:  MICHEL, MARIE-LOUISE
;  TITLE OF INVENTION: DESIGN OF A POLYEPITOPIC CONSTRUCT FOR THE INDUCTION
;  TITLE OF INVENTION:  OF
;  TITLE OF INVENTION:  HLA-A2.1 RESTRICTED HIV 1 SPECIFIC CTL RESPONSES USING
;  TITLE OF INVENTION:  HHD MICE
;  FILE REFERENCE: 03495.0196 SEQUENCE LISTING
;  CURRENT APPLICATION NUMBER: US/10/106,487
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: 09/675,673
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/158,356
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus type 1
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (9)
;   OTHER INFORMATION: L9V MUTANT EPITOPE
US-10-106-487-36

  Query Match             24.7%;  Score 49;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      23 LLWKGEGAV 31
           |||||||||
Db       1 LLWKGEGAV 9


RESULT 4
US-10-062-710-23
; Sequence 23, Application US/10062710
; Publication No. US20030049253A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Frank Q.
;  APPLICANT:  Chu, Yong-Liang
;  APPLICANT:  Qiu, Jian-Tai
;  TITLE OF INVENTION: Polymeric Conjugates for Delivery of
;  TITLE OF INVENTION:  MHC-Recognized Epitopes
;  TITLE OF INVENTION:  Via Peptide Vaccines
;  FILE REFERENCE: 3781-001-27
;  CURRENT APPLICATION NUMBER: US/10/062,710
;  CURRENT FILING DATE:  2002-02-05
;  PRIOR APPLICATION NUMBER: US 60/310,498
;  PRIOR FILING DATE: 2001-08-08
;  NUMBER OF SEQ ID NOS: 232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HIV Helper-T Cell Epitopes
US-10-062-710-23

  Query Match             24.7%;  Score 49;  DB 9;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYY 9
           |||||||||
Db       6 KIQNFRVYY 14


RESULT 5
US-09-764-868-871
; Sequence 871, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ32
;  CURRENT APPLICATION NUMBER: US/09/764,868
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1510
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 871
;   LENGTH: 96
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-868-871

  Query Match             23.5%;  Score 46.5;  DB 9;  Length 96;
  Best Local Similarity   34.2%;  Pred. No. 29;
  Matches   13;  Conservative    5;  Mismatches   11;  Indels    9;  Gaps    2;

Qy       4 NFRVYYRDSRDPLWKGPAK-------LLW--KGEGAVV 32
           || ::  |   | | | ::       |||  || | :|
Db      59 NFCIFSTDGVSPCWLGWSQFPDTVIHLLWPPKGLGLLV 96


RESULT 6
US-10-082-830-200
; Sequence 200, Application US/10082830
; Publication No. US20030077604A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Salceda, Susana
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Turner, Leah
;  TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
;  TITLE OF INVENTION:  Genes and Proteins
;  FILE REFERENCE: DEX-0249
;  CURRENT APPLICATION NUMBER: US/10/082,830
;  CURRENT FILING DATE:  2001-10-29
;  PRIOR APPLICATION NUMBER: 60/243,802
;  PRIOR FILING DATE: 2000-10-27
;  NUMBER OF SEQ ID NOS: 282
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 200
;   LENGTH: 38
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-082-830-200

  Query Match             23.2%;  Score 46;  DB 9;  Length 38;
  Best Local Similarity   43.8%;  Pred. No. 12;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 IQNFRVYYRDSRDPLW 17
           : |||:: ||   | |
Db       1 LANFRIFSRDRVSPCW 16


RESULT 7
US-10-125-540-534
; Sequence 534, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14C1
;  CURRENT APPLICATION NUMBER: US/10/125,540
;  CURRENT FILING DATE:  2002-04-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 534
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (75)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (76)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-534

  Query Match             23.0%;  Score 45.5;  DB 9;  Length 79;
  Best Local Similarity   28.2%;  Pred. No. 32;
  Matches   11;  Conservative    7;  Mismatches   10;  Indels   11;  Gaps    1;

Qy       4 NFRVYYRDS-----------RDPLWKGPAKLLWKGEGAV 31
           :|  ||:|:            | : :|||  : || | :
Db      33 HFSTYYKDNCPLHRKRAGHISDGIDQGPANSVSKGPGGL 71


RESULT 8
US-09-764-870-534
; Sequence 534, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14
;  CURRENT APPLICATION NUMBER: US/09/764,870
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 534
;   LENGTH: 79
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (28)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (75)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (76)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-534

  Query Match             23.0%;  Score 45.5;  DB 10;  Length 79;
  Best Local Similarity   28.2%;  Pred. No. 32;
  Matches   11;  Conservative    7;  Mismatches   10;  Indels   11;  Gaps    1;

Qy       4 NFRVYYRDS-----------RDPLWKGPAKLLWKGEGAV 31
           :|  ||:|:            | : :|||  : || | :
Db      33 HFSTYYKDNCPLHRKRAGHISDGIDQGPANSVSKGPGGL 71


RESULT 9
US-09-764-891-4912
; Sequence 4912, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4912
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-891-4912

  Query Match             22.7%;  Score 45;  DB 9;  Length 99;
  Best Local Similarity   36.4%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       2 IQNFRVYYRDSRDPLWKGPAKL 23
           |:. |       |:|| | |  :|
Db       6 IEQFLTIQEGSKDPFWAGRRRL 27


RESULT 10
US-09-764-891-5337
; Sequence 5337, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC006
;  CURRENT APPLICATION NUMBER: US/09/764,891
;  CURRENT FILING DATE:  2001-01-17
```

```
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 10231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5337
;    LENGTH: 93
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-764-891-5337

  Query Match             22.5%;  Score 44.5;  DB 9;  Length 93;
  Best Local Similarity   33.3%;  Pred. No. 52;
  Matches   10;  Conservative    3;  Mismatches   12;  Indels    5;  Gaps    1;

Qy       4 NFRVYYRDSRDPLWKGPAKLLWKGEGAVVI 33
           || :: ||   | | |      |     :||
Db      60 NFCIFSRDGVSPCWSG-----WSRTPDLVI 84


RESULT 11
US-09-764-877-1243
; Sequence 1243, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1243
;    LENGTH: 75
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: SITE
;    LOCATION: (56)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    NAME/KEY: SITE
;    LOCATION: (75)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1243

  Query Match             21.7%;  Score 43;  DB 10;  Length 75;
  Best Local Similarity   46.2%;  Pred. No. 67;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 YYRDSRDPLWKGP 20
           |: ::|  ||:||
Db      23 YHENTRAALWRGP 35


RESULT 12
US-09-989-920-243
; Sequence 243, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 243
;    LENGTH: 86
;    TYPE: PRT
;    ORGANISM: Homo sapien
US-09-989-920-243

  Query Match             21.7%;  Score 43;  DB 9;  Length 86;
  Best Local Similarity   50.0%;  Pred. No. 77;
  Matches    8;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 NFRVYYRDSRDPLWKG 19
           || |: ||   | | |
Db      49 NFCVFGRDGVSPCWPG 64


RESULT 13
US-10-011-585A-230
; Sequence 230, Application US/10011585A
; Publication No. US20030039986A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
;  TITLE OF INVENTION:  Genes and Proteins
;  FILE REFERENCE: DEX-0261
;  CURRENT APPLICATION NUMBER: US/10/011,585A
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: 60/245,740
;  PRIOR FILING DATE: 2000-11-03
;  NUMBER OF SEQ ID NOS: 245
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 230
;    LENGTH: 38
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (27)
;    OTHER INFORMATION: any amino acid
US-10-011-585A-230

  Query Match             20.7%;  Score 41;  DB 9;  Length 38;
  Best Local Similarity   41.9%;  Pred. No. 60;
  Matches   13;  Conservative    6;  Mismatches    6;  Indels    6;  Gaps    2;

Qy      10 RDSRDPLWKGPAKLL----WKGEGAVVIQDN 36
           ||||  | |||:: :     |||   |::|:
Db       9 RDSRRGLEKGPSRHMESQXQKGER--VLEDD 37


RESULT 14
US-09-864-761-36417
; Sequence 36417, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
```

```
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36417
;   LENGTH: 53
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007322.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.6
;   OTHER INFORMATION: EST_HUMAN HIT: AA196978.1, EVALUE 3.30e-01
US-09-864-761-36417

  Query Match             20.7%;  Score 41;  DB 10;  Length 53;
  Best Local Similarity   58.3%;  Pred. No. 86;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      25 WKGEGAVVIQDN 36
           |||  ||:| |:
Db       1 WKGSSAVLIADH 12


RESULT 15
US-09-734-569-120
; Sequence 120, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
;  APPLICANT: Lerchl, Jens
;  APPLICANT:  Renz, Andreas
;  APPLICANT:  Ehrhardt, Thomas
;  APPLICANT:  Reindl, Andreas
;  APPLICANT:  Cirpus, Petra
;  APPLICANT:  Bischoff, Friedrich
;  APPLICANT:  Frank, Markus
;  APPLICANT:  Freund, Annette
;  APPLICANT:  Duwenig, Elke
;  APPLICANT:  Schmidt, Ralf-Michael
;  APPLICANT:  Reski, Ralf
;  TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
;  TITLE OF INVENTION:  in the synthesis of carbohydrates
;  FILE REFERENCE: BASF-NAE-1332-99-US
;  CURRENT APPLICATION NUMBER: US/09/734,569
;  CURRENT FILING DATE:  2001-05-24
;  PRIOR APPLICATION NUMBER: US 60/171,101
;  PRIOR FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 120
;   LENGTH: 94
;   TYPE: PRT
;   ORGANISM: Physcomitrella patens
US-09-734-569-120

  Query Match             20.7%;  Score 41;  DB 10;  Length 94;
  Best Local Similarity   40.6%;  Pred. No. 1.6e+02;
  Matches   13;  Conservative    4;  Mismatches   11;  Indels    4;  Gaps    1;

Qy       7 VYYRDSRDP----LWKGPAKLLWKGEGAVVIQ 34
           | |  | |     | |  || |:|| || :::
Db      35 VKYNGSMDAFKQILAKEGAKSLFKGAGANILR 66


Search completed: May  6, 2003, 14:59:23
Job time : 17 secs
```

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:56:49 ; Search time 15 Seconds
(without alignments)
70.615 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 196705

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	79	39.9	15	4	US-09-009-953-221	Sequence 221, App
2	58	29.3	43	2	US-08-488-161-67	Sequence 67, Appl
3	58	29.3	43	3	US-09-273-685-67	Sequence 67, Appl
4	58	29.3	43	5	PCT-US95-11934-67	Sequence 67, Appl
5	54	27.3	10	3	US-08-159-339A-550	Sequence 550, App
6	54	27.3	15	4	US-09-009-953-220	Sequence 220, App
7	53	26.8	9	2	US-08-986-234-71	Sequence 71, Appl
8	53	26.8	10	4	US-08-197-484-74	Sequence 74, Appl
9	53	26.8	10	5	PCT-US95-02121-74	Sequence 74, Appl
10	49	24.7	9	2	US-08-986-234-72	Sequence 72, Appl
11	49	24.7	9	4	US-08-197-484-75	Sequence 75, Appl
12	49	24.7	9	5	PCT-US95-02121-75	Sequence 75, Appl
13	49	24.7	11	3	US-08-159-339A-1140	Sequence 1140, Ap
14	48	24.2	15	4	US-09-255-502-6	Sequence 6, Appli
15	46	23.2	43	2	US-08-488-161-51	Sequence 51, Appl
16	46	23.2	43	3	US-09-273-685-51	Sequence 51, Appl
17	46	23.2	43	5	PCT-US95-11934-51	Sequence 51, Appl
18	44	22.2	30	4	US-09-315-304B-1521	Sequence 1521, Ap
19	44	22.2	67	3	US-08-475-668A-210	Sequence 210, App
20	44	22.2	67	3	US-08-485-551A-210	Sequence 210, App
21	44	22.2	77	3	US-08-475-668A-211	Sequence 211, App
22	44	22.2	77	3	US-08-485-551A-211	Sequence 211, App
23	43	21.7	34	1	US-08-257-528B-26	Sequence 26, Appl
24	43	21.7	34	1	US-08-460-602A-26	Sequence 26, Appl
25	43	21.7	34	1	US-08-463-966A-26	Sequence 26, Appl
26	43	21.7	34	1	US-08-465-217A-26	Sequence 26, Appl
27	43	21.7	34	2	US-08-464-329A-26	Sequence 26, Appl
28	43	21.7	34	2	US-08-462-507A-26	Sequence 26, Appl
29	43	21.7	34	2	US-08-467-881A-26	Sequence 26, Appl
30	43	21.7	34	4	US-08-750-624-4	Sequence 4, Appli
31	43	21.7	37	1	US-08-257-528B-39	Sequence 39, Appl
32	43	21.7	37	1	US-08-460-602A-39	Sequence 39, Appl
33	43	21.7	37	1	US-08-463-966A-39	Sequence 39, Appl
34	43	21.7	37	1	US-08-465-217A-39	Sequence 39, Appl
35	43	21.7	37	2	US-08-464-329A-39	Sequence 39, Appl
36	43	21.7	37	2	US-08-462-507A-39	Sequence 39, Appl
37	43	21.7	37	2	US-08-467-881A-39	Sequence 39, Appl
38	43	21.7	40	1	US-08-426-819A-19	Sequence 19, Appl
39	41.5	21.0	66	4	US-09-082-593-6	Sequence 6, Appli
40	41	20.7	44	1	US-08-257-528B-95	Sequence 95, Appl
41	41	20.7	44	1	US-08-460-602A-95	Sequence 95, Appl
42	41	20.7	44	1	US-08-463-966A-95	Sequence 95, Appl
43	41	20.7	44	1	US-08-465-217A-95	Sequence 95, Appl
44	41	20.7	44	2	US-08-464-329A-95	Sequence 95, Appl
45	41	20.7	44	2	US-08-462-507A-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1
US-09-009-953-221
; Sequence 221, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
; Reactive DR Restricted Epitopes
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,953
; FILING DATE: 21-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 221:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 221:
US-09-009-953-221

Query Match 39.9%; Score 79; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;

```
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      22 KLLWKGEGAVVIQDN 36
           |||||||||||||||
Db       1 KLLWKGEGAVVIQDN 15


RESULT 2
US-08-488-161-67
; Sequence 67, Application US/08488161
; Patent No. 5885577
;  GENERAL INFORMATION:
;    APPLICANT:  Alvarez, Vernon L.
;    TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
;    TITLE OF INVENTION:  Peptide Libraries
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,161
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-176
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-488-161-67

  Query Match             29.3%;  Score 58;  DB 2;  Length 43;
  Best Local Similarity   66.7%;  Pred. No. 0.083;
  Matches   10;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      15 PLWKGPAKLLWKGEG 29
           | | ||| |||:| |
Db       9 PPWSGPAGLLWQGCG 23


RESULT 3
US-09-273-685-67
; Sequence 67, Application US/09273685
; Patent No. 6015561
;  GENERAL INFORMATION:
;    APPLICANT:  Alvarez, Vernon L.
;    TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
;    TITLE OF INVENTION:  Peptide Libraries
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/273,685
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/488,161
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-176
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-273-685-67

  Query Match             29.3%;  Score 58;  DB 3;  Length 43;
  Best Local Similarity   66.7%;  Pred. No. 0.083;
  Matches   10;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      15 PLWKGPAKLLWKGEG 29
           | | ||| |||:| |
Db       9 PPWSGPAGLLWQGCG 23


RESULT 4
PCT-US95-11934-67
; Sequence 67, Application PC/TUS9511934
;  GENERAL INFORMATION:
;    APPLICANT:  Cytogen Corporation
;    TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
;    TITLE OF INVENTION:  Peptide Libraries
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/11934
;      FILING DATE:  20-SEP-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-196-228
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  43 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US95-11934-67

  Query Match             29.3%;  Score 58;  DB 5;  Length 43;
  Best Local Similarity   66.7%;  Pred. No. 0.083;
  Matches   10;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      15 PLWKGPAKLLWKGEG 29
           | | ||| |||:| |
Db       9 PPWSGPAGLLWQGCG 23


RESULT 5
US-08-159-339A-550
; Sequence 550, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  550:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-550

  Query Match             27.3%;  Score 54;  DB 3;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.058;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYR 10
           ||||||||||
Db       1 KIQNFRVYYR 10


RESULT 6
US-09-009-953-220
; Sequence 220, Application US/09009953
; Patent No. 6413517
;  GENERAL INFORMATION:
;       APPLICANT: Sette, Alessandro
;       TITLE OF INVENTION: Identification of Broadly
;                           Reactive DR Restricted Epitopes
;       NUMBER OF SEQUENCES: 274
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/009,953
;            FILING DATE: 21-Jan-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 60/036,713
;            FILING DATE: 23-JAN-1997
;            APPLICATION NUMBER: US 60/037,432
;            FILING DATE: 07-FEB-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Weber, Ellen Lauver
;            REGISTRATION NUMBER: 32,762
;            REFERENCE/DOCKET NUMBER: 018623-011520US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-576-0200
;            TELEFAX: 415-576-0300
;            TELEX: <Unknown>
;  INFORMATION FOR SEQ ID NO: 220:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 15 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 220:
US-09-009-953-220

  Query Match             27.3%;  Score 54;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.094;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYYR 10
           ||||||||||
Db       6 KIQNFRVYYR 15


RESULT 7
US-08-986-234-71
; Sequence 71, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
;  APPLICANT: Wallen, et al.
;  TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
;  FILE REFERENCE: UNME-0008-1
;  CURRENT APPLICATION NUMBER: US/08/986,234
;  CURRENT FILING DATE: 1997-12-05
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 71
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-08-986-234-71

Query Match 26.8%; Score 53; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 PLWKGPAKL 23
||||||||| |
Db 1 PLWKGPAKL 9

RESULT 8
US-08-197-484-74
; Sequence 74, Application US/08197484
; Patent No. 6419931
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-197-484-74

Query Match 26.8%; Score 53; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 LLWKGEGAVV 32
||||||||| |
Db 1 LLWKGEGAVV 10

RESULT 9
PCT-US95-02121-74
; Sequence 74, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-74

Query Match 26.8%; Score 53; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 LLWKGEGAVV 32
||||||||| |
Db 1 LLWKGEGAVV 10

RESULT 10
US-08-986-234-72
; Sequence 72, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Wallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNME-0008-1

```
;  CURRENT APPLICATION NUMBER: US/08/986,234
;  CURRENT FILING DATE: 1997-12-05
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus
US-08-986-234-72

  Query Match             24.7%;  Score 49;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      23 LLWKGEGAV 31
           |||||||||
Db       1 LLWKGEGAV 9


RESULT 11
US-08-197-484-75
; Sequence 75, Application US/08197484
; Patent No. 6419931
;  GENERAL INFORMATION:
;    APPLICANT:  VITIELLO, Maria A.
;    APPLICANT:  CHESTNUT, Robert W.
;    APPLICANT:  SETTE, Alessandro D.
;    APPLICANT:  CELIS, Esteban
;    APPLICANT:  GRAY, Howard
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR ELICITING
;    TITLE OF INVENTION:  CTL IMMUNITY
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/197,484
;      FILING DATE:  16-FEB-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/935,811
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/874,491
;      FILING DATE:  27-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,682
;      FILING DATE:  29-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,568
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W.
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  14137-26-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 467-9600
;      TELEFAX:  (206) 623-6793
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-197-484-75

  Query Match             24.7%;  Score 49;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      23 LLWKGEGAV 31
           |||||||||
Db       1 LLWKGEGAV 9


RESULT 12
PCT-US95-02121-75
; Sequence 75, Application PC/TUS9502121
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR ELICITING
;    TITLE OF INVENTION:  CTL IMMUNITY
;    NUMBER OF SEQUENCES:  153
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02121
;      FILING DATE:  16-FEB-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/197,484
;      FILING DATE:  16-FEB-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/935,811
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/874,491
;      FILING DATE:  27-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,682
;      FILING DATE:  29-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,568
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W.
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  14137-26-4PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 467-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-02121-75

  Query Match             24.7%;  Score 49;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      23 LLWKGEGAV 31
           |||||||||
Db       1 LLWKGEGAV 9


RESULT 13
US-08-159-339A-1140
; Sequence 1140, Application US/08159339A
; Patent No. 6037135
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-1140

  Query Match             24.7%;  Score 49;  DB 3;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.36;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 KIQNFRVYY 9
            |||||||||
Db       3 KIQNFRVYY 11


RESULT 14
US-09-255-502-6
; Sequence 6, Application US/09255502
; Patent No. 6218165
; GENERAL INFORMATION:
;  APPLICANT:  Estell, David
;  APPLICANT:  Harding, Fiona
;  TITLE OF INVENTION:  Mutant Proteins Having Lower Allergenic Responses in
;  TITLE OF INVENTION:  Humans and Methods for Constructing, Identifying and
;  TITLE OF INVENTION:  Producing Such Proteins
;  FILE REFERENCE:  GC 527-D2
;  CURRENT APPLICATION NUMBER:  US/09/255,502
;  CURRENT FILING DATE:  1999-02-23
;  PRIOR APPLICATION NUMBER:  09/060,872
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;    LENGTH:  15
;    TYPE:  PRT
;    ORGANISM:  Unknown
;    FEATURE:
;    OTHER INFORMATION:  Description of Unknown Organism:  Unknown Sequence
US-09-255-502-6

  Query Match             24.2%;  Score 48;  DB 4;  Length 15;
  Best Local Similarity   61.5%;  Pred. No. 0.72;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 IQNFRVYYRDSRD 14
            |::| ||:|:|||
Db       1 IKDFHVYFRESRD 13


RESULT 15
US-08-488-161-51
; Sequence 51, Application US/08488161
; Patent No. 5885577
;  GENERAL INFORMATION:
;    APPLICANT:  Alvarez, Vernon L.
;    TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
;    TITLE OF INVENTION:  Peptide Libraries
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,161
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-176
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-488-161-51

  Query Match             23.2%;  Score 46;  DB 2;  Length 43;
  Best Local Similarity   43.8%;  Pred. No. 4.9;
  Matches    7;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      14 DPLWKGPAKLLWKGEG 29
             :| |:||    | | |
Db       8 NPRWRGPGGFKWPGCG 23


Search completed:  May  6, 2003, 14:59:00
Job time :  16 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:54:59 ; Search time 34 Seconds
(without alignments)
141.089 Million cell updates/sec

Title: US-09-868-399-1
Perfect score: 198
Sequence: 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 575469

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	198	100.0	36	21	AAB15399	HIV integrase pept
2	169	85.4	31	9	AAP82730	Immunologically mi
3	169	85.4	50	9	AAP82728	Immunologically mi
4	161	81.3	30	23	AAU84494	HIV POL segment 63
5	110	55.6	30	23	AAU84493	HIV POL segment 62
6	90	45.5	15	22	ABP24758	HIV DR super motif
7	88	44.4	15	22	ABP24701	HIV DR super motif
8	88	44.4	15	22	ABP25020	HIV DR 3b motif po
9	86	43.4	15	22	ABP24677	HIV DR super motif
10	86	43.4	15	22	ABP25016	HIV DR 3b motif po
11	83	41.9	30	23	AAU84495	HIV POL segment 64
12	79	39.9	15	19	AAW85346	Helper T-cell clas
13	79	39.9	15	22	ABP24643	HIV DR super motif
14	70	35.4	15	11	AAR03899	HIV-antibody react
15	69	34.8	15	22	ABP24663	HIV DR super motif
16	68	34.3	11	22	ABP19154	HIV B62 super moti
17	66	33.3	11	22	ABP19309	HIV B62 super moti
18	62	31.3	10	22	ABP16223	HIV A24 super moti
19	62	31.3	10	22	ABP24215	HIV A24 motif pol
20	61	30.8	10	22	ABP16227	HIV A24 super moti
21	61	30.8	10	22	ABP21162	HIV A03 motif pol
22	61	30.8	10	22	ABP24230	HIV A24 motif pol
23	61	30.8	11	22	ABP13975	HIV A02 super moti
24	61	30.8	11	22	ABP16925	HIV B07 super moti
25	60	30.3	10	22	ABP16222	HIV A24 super moti
26	60	30.3	10	22	ABP21190	HIV A03 motif pol
27	60	30.3	10	22	ABP23194	HIV A11 motif pol
28	60	30.3	10	22	ABP24264	HIV A24 motif pol
29	60	30.3	11	22	ABP13974	HIV A02 super moti
30	59	29.8	10	22	ABP16226	HIV A24 super moti
31	59	29.8	10	22	ABP16893	HIV B07 super moti
32	59	29.8	10	22	ABP16926	HIV B07 super moti
33	59	29.8	10	22	ABP21518	HIV A03 motif pol
34	59	29.8	10	22	ABP24268	HIV A24 motif pol
35	58	29.3	10	22	ABP21515	HIV A03 motif pol
36	58	29.3	10	22	ABP23424	HIV A11 motif pol
37	58	29.3	11	22	ABP13973	HIV A02 super moti
38	58	29.3	11	22	ABP13976	HIV A02 super moti
39	58	29.3	11	22	ABP16382	HIV A24 super moti
40	58	29.3	11	22	ABP19302	HIV B62 super moti
41	58	29.3	11	22	ABP19464	HIV B62 super moti
42	58	29.3	43	17	AAR95500	E88, monoclonal an
43	57	28.8	9	22	ABP17386	HIV B27 super moti
44	57	28.8	10	22	ABP11884	HIV A01 super moti
45	57	28.8	10	22	ABP13740	HIV A02 super moti

ALIGNMENTS

RESULT 1
AAB15399
ID AAB15399 standard; peptide; 36 AA.
XX
AC AAB15399;
XX
DT 15-DEC-2000 (first entry)
XX
DE HIV integrase peptide epitope.
XX
KW Detection; infection; HIV-1; antigen; pol; mixotope; epitope; integrase;
KW combinatorial peptide; antibody.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200040970-A1.
XX
PD 13-JUL-2000.
XX
PF 29-DEC-1999; 99WO-FR03311.
XX
PR 31-DEC-1998; 98FR-0016727.
XX
PA (INSP) INST PASTEUR LILLE.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (GRAS/) GRAS-MASSE H.
XX
PI Tranchand-Bunel D, Auriault C;
XX
DR WPI; 2000-475858/41.
XX
PT Reagent for detecting human immune deficiency virus infection, useful
PT for diagnosis and monitoring, comprises pol-derived antigen and mixture

PT of its convergent peptides -
XX
PS Claim 3; Page 15; 32pp; French.
XX
CC The invention relates to a reagent for detecting infection by human
CC immunodeficiency virus (HIV), comprising a mixture of an antigenic
CC peptide encoded by the pol gene and containing at most 60 (preferably
CC 20-40) amino acids, and a mixture, designated a mixotope, of converged,
CC combinatorial peptides derived from the antigenic peptide. This sequence
CC corresponds to an epitope from the integrase protein encoded by the pol
CC gene from HIV-1. The reagent is used to detect anti-HIV antibodies,
CC specifically for diagnosis and monitoring of infection.
XX
SQ Sequence 36 AA;

Query Match 100.0%; Score 198; DB 21; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36
||||||||||||||||||||||||||||||||||||
Db 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDN 36

RESULT 2
AAP82730
ID AAP82730 standard; protein; 31 AA.
XX
AC AAP82730;
XX
DT 23-NOV-1990 (first entry)
XX
DE Immunologically mimicking peptide.
XX
KW Immunological mimicry; HIV gag and pol; AIDS.
XX
OS synthetic.
XX
PN EP267802-A.
XX
PD 18-MAY-1988.
XX
PF 13-NOV-1987; 87EP-0310047.
XX
PR 14-NOV-1986; 86US-0930785.
XX
PA (GENE-) GENETIC SYST CORP.
XX
PI Cosand WL, Harris LJ, Houghton RL;
XX
DR WPI; 1988-134645/20.
DR N-PSDB; AAN82239.
XX
PT Synthetic peptide(s) having sequence corresp. to HIV virus - used
PT for detection of AIDS-related disease and in vaccine prepn.
XX
PS Claim 14; Page 12; 12pp; English.
XX
CC This peptide immunologically mimics a protein encoded by the HIV
CC retrovirus and corresps. to part of peptide 123. It is competitive
CC with HIV and complexes formed on binding of antibodies are detected.
CC It is therefore used as an accurate test for detecting patients who
CC have been exposed to the aetiological agent of lymphadenopathy
CC syndrome and/or AIDS. Residue 31 can be absent but is opt. present.
CC See also AAN82236-38.
XX
SQ Sequence 31 AA;

Query Match 85.4%; Score 169; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 5e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGA 30
||||||||||||||||||||||||||||||
Db 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGA 30

RESULT 3
AAP82728
ID AAP82728 standard; protein; 50 AA.
XX
AC AAP82728;
XX
DT 23-NOV-1990 (first entry)
XX
DE Immunologically mimicking peptide.
XX
KW Immunological mimicry; HIV gag and pol; AIDS.
XX
OS synthetic.
XX
PN EP267802-A.
XX
PD 18-MAY-1988.
XX
PF 13-NOV-1987; 87EP-0310047.
XX
PR 14-NOV-1986; 86US-0930785.
XX
PA (GENE-) GENETIC SYST CORP.
XX
PI Cosand WL, Harris LJ, Houghton RL;
XX
DR WPI; 1988-134645/20.
DR N-PSDB; AAN82237.
XX
PT Synthetic peptide(s) having sequence corresp. to HIV virus - used
PT for detection of AIDS-related disease and in vaccine prepn.
XX
PS Claim 14; Page 12; 12pp; English.
XX
CC This peptide which immunologically mimics a protein encoded by the
CC HIV retrovirus corresps. to amino acids 4385 to 4519 of pol
CC protein p31. It is competitive with HIV and complexes formed on
CC binding of antibodies are detected. It is therefore used as an
CC accurate test for detecting patients who have been exposed to the
CC aetiological agent of lymphadenopathy syndrome and/or AIDS.
CC Residues 1-4 and 50 can be absent but are opt. present.
CC See also AAN82236 and AAN82238-39.
XX
SQ Sequence 50 AA;

Query Match 85.4%; Score 169; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 8.9e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGA 30
||||||||||||||||||||||||||||||
Db 20 KIQNFRVYYRDSRDPLWKGPAKLLWKGEGA 49

RESULT 4
AAU84494
ID AAU84494 standard; Peptide; 30 AA.
XX
AC AAU84494;
XX
DT 08-MAY-2002 (first entry)
XX
DE HIV POL segment 63.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficieny virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX

OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU00622.
XX
PR 26-MAY-2000; 2000AU-0007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36333.
XX
PT New synthetic polypeptides having several different segments of at
PT least one parent polypeptide linked together differently compared to
PT the linkage in the parent polypeptide, for inducing immune response
PT against a pathogen or cancer -
XX
PS Example 1; Fig 12; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for
CC designing the synthetic polypeptides. The synthetic polypeptides and
CC polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC useful for modulating immune responses preferably directed against a
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present
CC sequence is a peptide derived from a parent protein used to
CC construct a savine of the invention.
XX
SQ Sequence 30 AA;

Query Match 81.3%; Score 161; DB 23; Length 30;
Best Local Similarity 96.7%; Pred. No. 7.5e-17;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 RVYYRDSRDPLWKGPAKLLWKGEGAVVIQD 35
||||||||| ||||||||||||||||||||
Db 1 RVYYRDSRDPXWKGPAKLLWKGEGAVVIQD 30

RESULT 5
AAU84493
ID AAU84493 standard; Peptide; 30 AA.
XX
AC AAU84493;
XX
DT 08-MAY-2002 (first entry)
XX
DE HIV POL segment 62.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficieny virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU00622.
XX
PR 26-MAY-2000; 2000AU-0007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36332.
XX
PT New synthetic polypeptides having several different segments of at
PT least one parent polypeptide linked together differently compared to
PT the linkage in the parent polypeptide, for inducing immune response
PT against a pathogen or cancer -
XX
PS Example 1; Fig 12; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for
CC designing the synthetic polypeptides. The synthetic polypeptides and
CC polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC useful for modulating immune responses preferably directed against a
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present
CC sequence is a peptide derived from a parent protein used to
CC construct a savine of the invention.
XX
SQ Sequence 30 AA;

Query Match 55.6%; Score 110; DB 23; Length 30;
Best Local Similarity 95.0%; Pred. No. 3.1e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSRDPLWKGP 20
|||||||||||||||| ||||
Db 11 KIQNFRVYYRDSRDPXWKGP 30

RESULT 6
ABP24758
ID ABP24758 standard; Peptide; 15 AA.
XX
AC ABP24758;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif pol peptide #125.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;

KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 376; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 45.5%; Score 90; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 RDPLWKGPAKLLWKG 27
||||||||||||||||
Db 1 RDPLWKGPAKLLWKG 15

RESULT 7
ABP24701
ID ABP24701 standard; Peptide; 15 AA.
XX
AC ABP24701;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif pol peptide #68.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 375; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 44.4%; Score 88; DB 22; Length 15;
Best Local Similarity 93.3%; Pred. No. 2.6e-06;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 RDPLWKGPAKLLWKG 27
|||:|||||||||||
Db 1 RDPIWKGPAKLLWKG 15

RESULT 8
ABP25020
ID ABP25020 standard; Peptide; 15 AA.
XX
AC ABP25020;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR 3b motif pol peptide #11.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.

XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 408; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 44.4%; Score 88; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FRVYYRDSRDPLWKG 19
||||||||||||||||
Db 1 FRVYYRDSRDPLWKG 15

RESULT 9
ABP24677
ID ABP24677 standard; Peptide; 15 AA.
XX
AC ABP24677;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif pol peptide #44.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 374; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 43.4%; Score 86; DB 22; Length 15;
Best Local Similarity 93.3%; Pred. No. 5.2e-06;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 NFRVYYRDSRDPLWK 18
|||||||||||||:||
Db 1 NFRVYYRDSRDPIWK 15

RESULT 10
ABP25016
ID ABP25016 standard; Peptide; 15 AA.
XX
AC ABP25016;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR 3b motif pol peptide #7.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.

XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 408; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 43.4%; Score 86; DB 22; Length 15;
Best Local Similarity 93.3%; Pred. No. 5.2e-06;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FRVYYRDSRDPLWKG 19
||||||||||||:|||
Db 1 FRVYYRDSRDPIWKG 15

RESULT 11
AAU84495
ID AAU84495 standard; Peptide; 30 AA.
XX
AC AAU84495;
XX
DT 08-MAY-2002 (first entry)
XX
DE HIV POL segment 64.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU00622.
XX
PR 26-MAY-2000; 2000AU-0007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI; 2002-147575/19.
DR N-PSDB; ABK36334.
XX
PT New synthetic polypeptides having several different segments of at
PT least one parent polypeptide linked together differently compared to
PT the linkage in the parent polypeptide, for inducing immune response
PT against a pathogen or cancer -
XX
PS Example 1; Fig 12; 364pp; English.
XX
CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for
CC designing the synthetic polypeptides. The synthetic polypeptides and
CC polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC useful for modulating immune responses preferably directed against a
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present
CC sequence is a peptide derived from a parent protein used to
CC construct a savine of the invention.
XX
SQ Sequence 30 AA;

Query Match 41.9%; Score 83; DB 23; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 AKLLWKGEGAVVIQDN 36
||||||||||||||||
Db 1 AKLLWKGEGAVVIQDN 16

RESULT 12
AAW85346
ID AAW85346 standard; peptide; 15 AA.
XX
AC AAW85346;
XX
DT 16-FEB-1999 (first entry)
XX
DE Helper T-cell class II peptide derived from POL protein.
XX
KW Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1;
KW DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW acquired immune deficiency syndrome; malaria; cancer;
KW allograft rejection; allergy; Lyme disease; hepatitis;
KW post-streptococcal endocarditis; glomerulonephritis;
KW food hypersensitivity.
XX
OS Synthetic.
OS Human immunodeficiency virus type 1.
XX
PN WO9832456-A1.
XX
PD 30-JUL-1998.
XX

PF 23-JAN-1998; 98WO-US01373.
XX
PR 07-FEB-1997; 97US-0037432.
PR 23-JAN-1997; 97US-0036713.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S;
XX
DR WPI; 1998-427679/36.
XX
PT Composition containing peptide that induces cytotoxic T lymphocyte
PT response, and helper peptide - can bind to human leucocyte antigen
PT alleles, used to treat or prevent cancers, parasitic infections and
PT autoimmune disease
XX
PS Disclosure; Page 41; 51pp; English.
XX
CC AAW85284-451 represent helper T-cell class II peptides, which can bind
CC to the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC are used in the course of the invention. The specification describes
CC peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC T-helper peptides, that are used together to generate a CTL response for
CC the treatment or prevention of viral, fungal, bacterial or parasitic
CC infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC alone to induce a helper T cell response, e.g. in cases of autoimmune
CC disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC post-streptococcal endocarditis, glomerulonephritis and food
CC hypersensitivity.
XX
SQ Sequence 15 AA;

Query Match 39.9%; Score 79; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 KLLWKGEGAVVIQDN 36
||||||||||||||||
Db 1 KLLWKGEGAVVIQDN 15

RESULT 13
ABP24643
ID ABP24643 standard; Peptide; 15 AA.
XX
AC ABP24643;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif pol peptide #10.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 374; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 39.9%; Score 79; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 KLLWKGEGAVVIQDN 36
||||||||||||||||
Db 1 KLLWKGEGAVVIQDN 15

RESULT 14
AAR03899
ID AAR03899 standard; peptide; 15 AA.
XX
AC AAR03899;
XX
DT 16-FEB-1993 (first entry)
XX
DE HIV-antibody reactive peptide (2).
XX
KW HIV; diagnosis.
XX
OS Synthetic.
XX
PN EP362915-A.
XX
PD 11-APR-1990.
XX
PF 07-SEP-1989; 89EP-0202258.
XX
PR 09-SEP-1988; 88NL-0002217.
XX
PA (ALKU) AKZO NV.
XX
PI Hellings JA, Schalken JJ, Sprengers ED;
XX
DR WPI; 1990-109271/15.
XX
PT New synthetic oligopeptide cpds. reactive with HIV antibodies -
PT useful as diagnostic reagents are non infectious and safe to use
XX
PS Claim 3; Page 5 + Fig 2; 7pp; English.
XX

CC The peptides given in AAR03898-900 react immunochemically with
CC antibodies directed against HIV. They are suitable for use in a
CC diagnostic method for determining the presence of HIV or HIV-
CC antibodies in a test fluid. In contrast to the native HIV, the
CC peptides have the great advantage that these are of a safe non-
CC infectious origin. They are made by standard (esp. solid phase)
CC methods of peptide synthesis, or by recombinant DNA techniques.
XX
SQ Sequence 15 AA;

Query Match 35.4%; Score 70; DB 11; Length 15;
Best Local Similarity 92.9%; Pred. No. 0.0013;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSRD 14
||||||||||||:
Db 2 KIQNFRVYYRDSRN 15

RESULT 15
ABP24663
ID ABP24663 standard; Peptide; 15 AA.
XX
AC ABP24663;
XX
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif pol peptide #30.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW vpu; vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope;
KW antigen; vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US27766.
XX
PR 05-OCT-1999; 99US-0412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX
PS Claim 32; Page 374; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines.
CC An additional advantage of an group-based vaccine approach is the ability
CC to combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 15 AA;

Query Match 34.8%; Score 69; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQNFRVYYRDSR 13
|||||||||||||
Db 3 KIQNFRVYYRDSR 15

Search completed: May 6, 2003, 14:57:23
Job time : 35 secs